METHOD FOR ENHANCEMENT IN SCREENING THROUGHPUT

BACKGROUND OF THE INVENTION

[0001]   This invention relates to rapid throughput screening of combinatorial libraries. More particularly, the invention relates to methods and devices for reducing data acquisition time for spectroscopic analysis and imaging of multiple samples.

[0002]   From the analytical perspective, quantification of a chemical property requires that the data can be analyzed with high confidence that any variation in the measured signal is due to a change of a chemical property of the sample, and not to variation in the measurement system.   The quality of analysis of samples from combinatorial libraries is determined by a number of factors, the most important of which are the speed of measurement (screening throughput), reproducibility (precision), and accuracy. Precision generally will improve as the measured quantity due to the parameter of interest (signal) increases in comparison to the measured quantity due to general background (noise). For a variety of analytical techniques, increasing signal integration time improves signal quality. For high throughput screening of large numbers of samples, however, increasing integration time is often impractical.

[0003]   Also, in many cases, the measured signal is proportional to sample size. For example, in Raman and fluorescence spectroscopy, the measured signal is directly proportional to the illuminated sample volume. Similarly, in absorption, mass-spectrometric, ionization, gravimetric, and other types of detection systems, the sample must be large enough to provide a detectable change in instrument response. However, developing new products and processes often requires combinatorial approaches that utilize multiple small-scale reactions arranged in an array format. Although efficient from the standpoint of cost, the small sample size is often accompanied by a decrease in signal, thus requiring long data acquisition times per sample.

[0004]   Although many analytical techniques can be used to analyze libraries of small individual samples, these detection schemes are generally not practical because of the long sample integration times needed to analyze small samples. Thus, there is a need to develop a method by which signal integration time, and therefore overall data acquisition time, is reduced to a point where it no longer limits methods of analysis.

SUMMARY OF THE INVENTION

[0005]   The present invention comprises methods and devices for reducing sample integration time using mathematical transform analysis and, thereby, significantly reducing the time required for screening multiple samples.   The invention is particularly suited for screening combinatorial libraries comprising multiple samples of small volume.

[0006]   In one aspect, the invention comprises a method for reducing the time required for analyzing at least one sample for a parameter of interest which comprises collecting analytical data from a sample using a predetermined integration time T_a and applying mathematical transform analysis on the data, wherein the mathematical transform analysis is performed using conditions designed to achieve a pre-determined signal quality response function value comprising the value obtained when samples are analyzed without mathematical transform analysis using integration time T_b, wherein T_b is greater than T_a.

[0007]   In another aspect, the invention comprises an apparatus for analyzing at least one sample for a parameter of interest using a pre-determined signal quality response function which comprises: a collecting system for collecting analytical data comprising the parameter of interest from a sample; a processing system for processing the analytical data; a screening rate accelerator toolbox for applying mathematical transform analysis to the data; a data analysis system for determining whether the data processed by the screening rate accelerator toolbox satisfies a pre-determined signal quality response function value; and a statistical toolbox for analyzing the processed data for the parameter of interest.

[0008]   The foregoing focuses on the more important features of the invention in order that the detailed description which follows may be better understood and in order that the present contribution to the art may be appreciated. There are additional features of the invention which will be described hereinafter and which will form the subject matter of the claims appended hereto. It is to be understood that the invention is not limited in its application to the details as set form in the following description and drawings, but is capable of other embodiments and of being practiced or carried out in various ways.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009]   Various features, aspects and advantages of the present invention will become more apparent with reference to the following description, appended claims, and accompanying drawings, wherein:

[0010]   Figure 1 is a flow chart illustrating an embodiment of the overall method of the present invention;

[0011]   Figure 2 illustrates an aspect of an embodiment of the method of the invention detailing steps for determination of the minimal integration time that satisfies a signal quality response function;

[0012]   Figure 3 illustrates an aspect of an embodiment of the method of the invention detailing the application of the screening rate accelerator toolbox in serial or parallel screening of an array of samples;

[0013]   Figure 4 illustrates an aspect of an embodiment of the apparatus of the invention comprising a system for serial screening of combinatorial libraries;

[0014]   Figure 5 illustrates an aspect of an embodiment of the apparatus of the invention comprising a system for parallel screening of combinatorial libraries;

[0015]   Figure 6 illustrates an aspect of an embodiment of the invention showing an original spectrum (dotted line) and a wavelet processed spectrum (solid line) in the

region of the 1072 cm^{-1} band of a Raman spectrum of diphenyl carbonate (DPC) using the screening rate accelerator toolbox of the invention;

[0016]   Figure 7 illustrates an aspect of an embodiment of the invention showing an original spectrum (dotted line) and a wavelet processed spectrum (solid line) in the region of the 911 cm^{-1} band of a Raman spectrum of diphenyl carbonate (DPC) using the screening rate accelerator toolbox of the invention;

[0017]   Figure 8 illustrates an aspect of an embodiment of the invention comprising a reduction in integration time for Raman analysis of diphenyl carbonate (DPC) at the 1072 cm^{-1} band where solid circles are before, and open squares are after, signal processing by the method of the invention;

[0018]   Figure 9 illustrates an aspect of an embodiment of the invention comprising a reduction in integration time for Raman analysis of diphenyl carbonate (DPC) at the 911 cm^{-1} band where solid circles are before, and open squares are after, signal processing by the method of the invention;

[0019]   Figure 10 illustrates an aspect of an embodiment of the invention showing the reduction of integration time ΔT ($\Delta T = T_b - T_a$) at a predetermined signal-to-noise ratio (S/N) in a collected signal before (T_b) and after (T_a) applying the screening rate accelerator toolbox in Raman analysis of diphenyl carbonate (DPC) at the 1072 cm^{-1} band;

[0020]   Figure 11 illustrates an aspect of an embodiment of the invention showing the reduction of integration time ΔT ($\Delta T = T_b - T_a$) at a predetermined signal-to-noise ratio (S/N) in a collected signal before (T_b) and after (T_a) applying the screening rate accelerator toolbox in Raman analysis of diphenyl carbonate (DPC) at the 911 cm^{-1} band; and

[0021]   Figure 12 illustrates an aspect of an embodiment of the invention showing the improvement in screening time of combinatorial libraries of different numbers of individual samples wherein open circles comprise total analysis time for a given library of original Raman spectra before processing by the method of the invention

4

and closed circles comprise total analysis time for a given library of spectra after signal processing by the method of the invention.

DETAILED DESCRIPTION OF THE INVENTION

[0022] The present invention provides methods and an apparatus for the rapid analysis of data from imaging, spectroscopic, scanning probe, or sensor methods by the application of mathematical transform analysis to reduce sample integration time. For example, using wavelet transform analysis, sample integration time is reduced by more than 100-fold. This reduction in integration time results in significant savings in the time required for data analysis.

[0023] In one aspect, the invention comprises a method for reducing the time required for analyzing at least one sample for a parameter of interest which comprises collecting analytical data from a sample using a predetermined integration time T_a and applying mathematical transform analysis on the data, wherein the mathematical transform analysis is performed using conditions designed to achieve a pre-determined signal quality response function value comprising the value obtained when samples are analyzed without mathematical transform analysis using integration time T_b, wherein T_b is greater than T_a.

[0024] In an embodiment, the mathematical transform analysis comprises multivariate analysis. Preferably, the multivariate analysis comprises neural networks analysis, principal components analysis, partial least squares analysis, linear multivariate analysis, or nonlinear multivariate analysis.

[0025] In another embodiment, the mathematical transform analysis comprises discrete transform analysis. Alternatively, the mathematical transform analysis may comprise continuous transform analysis.

[0026] In yet another embodiment, the mathematical transform analysis comprises time averaging analysis, smoothing analysis or Savitsky-Golay analysis. The mathematical transform analysis may also comprise Fourier transform, Gabor transform, or Hadamard transform.

[0027]   In an embodiment, the mathematical transform analysis comprises wavelet transform. In an embodiment, the wavelet transform analysis comprises a wavelet de-noising algorithm. Preferably, the wavelet de-noising algorithm comprises wavelet filters.   Also preferably, the wavelet de-noising algorithm comprises a threshold/shrinkage method. Preferably, parameters of the mathematical transform are determined during the course of analysis.

[0028]   Preferably, the preset signal quality response function ($y = f(x_1, x_2, x_3, \ldots x_n)$), comprises one or more measured signal parameters, (x), either used individually or combined mathematically. In an embodiment, at least one of the measured signal parameters comprises a signal-to-noise ratio. In another embodiment, at least one of the measured signal parameters comprises signal resolution. In another embodiment, at least one of the measured signal parameters comprises peak shift. In yet another embodiment, at least one of the measured signal parameters comprises signal distortion.   Thus, the signal quality response function may consist of one, or a combination of signal parameters.

[0029]   Where at least one of the signal parameters comprises a signal-to-noise ratio, preferably, the pre-determined signal quality response function value comprises a signal to noise ratio which ranges from 1 to 10,000. More preferably, the signal to noise ratio ranges from 2 to 5,000. Even more preferably, the signal to noise ratio ranges from 3 to 1,000.

[0030]   In an embodiment, the analytical data comprises a first-order array. In an alternate embodiment, the analytical data comprises a multi-order array.

[0031]   Preferably, the analytical data comprise an increase in signal resolution as a function of acquisition time.   Also preferably, the analytical data comprise spectroscopic, imaging, sensor, or scanning data. More preferably, the data further comprise measurements made using Raman, luminescence, ultraviolet-visible molecular absorbance, atomic absorbance, infra-red, near infrared, surface plasmon resonance, mass spectrometry, X-ray, nuclear magnetic resonance, refractometry, interferometry, scattering, inductively coupled plasma, atomic force microscopy,

6

scanning tunneling microscopy, microwave evanescent wave microscopy, near-field scanning optical microscopy, atomic fluorescence, laser-induced breakdown spectroscopy, Auger electron spectroscopy, X-ray photoelectron spectroscopy, ultrasonic spectroscopy, dielectric spectroscopy, microwave spectroscopy, or resonance-enhanced multiphoton ionization, and the like. Also, combinations of these techniques can be used, for example surface plasmon resonance and fluorescence, Raman and infrared, and any others. Also, different improvements and subclasses of these techniques can be used, for example, resonance Raman, surface-enhanced Raman, resonance surface-enhanced Raman, time-of-flight mass spectrometry, secondary ion mass spectrometry, ion mobility spectrometry, and the like. The techniques used to collect the analytical data may also comprise photon probe microscopy, electron probe microscopy, ion probe microscopy, field probe microscopy, scanning probe microscopy, and the like. In an embodiment, analytical data is provided using techniques relying on collection of electromagnetic radiation in the range from 0.05 Angstroms to 500 millimeters (mm).

[0032] The sample may be inorganic material, organic material, polymeric material, biological material, or combinations thereof. In an embodiment, the sample comprises polycarbonate. Further, the concentration ranges of species of interest analyzed using these techniques can range from detected single molecules to concentrations of up to 100 percent of materials of interest. Thus, in an embodiment, the sample may comprise a single molecule. In another embodiment, the parameter of interest may comprise up to 100% of the sample. Any type of single molecule spectroscopy and imaging can utilize the method of this invention.

[0033] In an embodiment, the invention comprises computer readable media comprising software code for performing the methods of the invention. The present invention also comprises systems using the methods of the invention.

[0034] In another aspect, the invention comprises a method for reducing the time required for analyzing at least one sample for a parameter of interest which comprises: selecting a pre-determined signal quality response function; selecting a pre-determined

7

integration time T_a; collecting analytical data from the sample with integration time T_a; applying a screening'rate accelerator toolbox comprising mathematical transform analysis to the data, wherein the mathematical transform analysis is performed using conditions designed to achieve a pre-determined signal quality response function value comprising the value obtained when samples are analyzed without mathematical transform analysis using integration time T_b, wherein T_b is greater than T_a; and analyzing the data processed by the screening rate accelerator toolbox for the parameter of interest.

[0035] In an embodiment, the mathematical transform analysis comprises multivariate analysis. Preferably, the multivariate analysis comprises neural networks analysis, principal components analysis, partial least squares analysis, linear multivariate analysis, or nonlinear multivariate analysis.

[0036] In another embodiment, the mathematical transform analysis comprises discrete transform analysis. Alternatively, the mathematical transform analysis may comprise continuous transform analysis.

[0037] In yet another embodiment, the mathematical transform analysis comprises time averaging analysis, smoothing analysis or Savitsky-Golay analysis. The mathematical transform analysis may also comprise Fourier transform, Gabor transform, or Hadamard transform.

[0038] In an embodiment, the mathematical transform analysis comprises wavelet transform. In an embodiment, the wavelet transform analysis comprises a wavelet de-noising algorithm. Preferably, the wavelet de-noising algorithm comprises wavelet filters. Also preferably, the wavelet de-noising algorithm comprises a threshold/shrinkage method. Preferably, parameters of the mathematical transform are determined during the course of analysis.

[0039] The method allows for flexibility in the decision as to whether the screening rate accelerator toolbox is required. In an embodiment, the method comprises the steps of determining whether the analytical data collected from the sample with

integration time T_a satisfies the pre-determined signal quality response function value prior to applying the screening rate accelerator toolbox and applying the screening rate accelerator toolbox if the collected data does not satisfy the pre-determined signal quality response function value, but not if the collected data does satisfy the pre-determined signal quality response function value.

[0040] The method also permits iterative application of the screening rate accelerator toolbox. Thus, in an embodiment, the method includes the steps of: determining whether the data processed by the screening rate accelerator toolbox satisfies the pre-determined signal quality response function value; and if the processed data does not satisfy the pre-determined signal quality response value, re-applying the screening rate accelerator toolbox using a mathematical transform analysis different from the analysis previously applied until the data processed using the screen rate accelerator toolbox either satisfies the pre-determined signal quality response function value or comprises an optimized signal quality response function value.

[0041] In some cases, even with application of the screening rate accelerator toolbox, the data collected with an integration time T_a may not satisfy the signal quality response function value. Thus, in an embodiment, the method includes the steps of repeating the method with a new, larger value for T_a if the data which is optimized using the screening rate accelerator toolbox does not satisfy the pre-determined signal quality response function value.

[0042] One purpose of the method is to minimize the time required for processing large arrays of experimental samples. Thus, in an embodiment, if the collected analytical data does not require application of the screening rate accelerator toolbox to satisfy a pre-determined signal quality response function value, a shorter pre-determined integration time T_a is selected, and the method performed using the new value for T_a. By reducing T_a to a value which allows for the pre-determined signal quality response function value (e.g. S/N = 20) to be obtained upon application of the

screening rate accelerator toolbox, the overall time required for sample analysis is minimized.

[0043] Preferably, the preset signal quality response function comprises one or more measured signal parameters, either used individually or combined mathematically. In an embodiment, at least one of the measured signal parameters comprises a signal-to-noise ratio. Preferably, the signal quality response function value comprises a signal to noise ratio of 1 to about 10,000. More preferably, the signal quality response function value comprises a signal to noise ratio of 2 to 5,000. Even more preferably, the signal quality response function value comprises a signal to noise ratio of 3 to 1,000.

[0044] In another embodiment, at least one of the measured signal parameters comprises signal resolution. In another embodiment, at least one of measured the signal parameters comprises peak shift. In yet another embodiment, at least one of the signal parameters comprises signal distortion.

[0045] In an embodiment, the relative improvement in signal integration time (T_b/T_a) ranges from about 1.5 to 1,000 fold. In another embodiment, the relative improvement in signal integration time (T_b/T_a) ranges from about 1.5 to 500 fold. In yet another embodiment, the relative improvement in signal integration time (T_b/T_a) ranges from about 1.5 to 200 fold.

[0046] In an embodiment, the analytical data comprises a first-order array. In an alternate embodiment, the analytical data comprises a second-order or higher array such as multi-order arrays and the like.

[0047] In an embodiment, evaluation of each individual sample in an array of samples is substantially simultaneous. In an alternate embodiment, each individual sample in an array of samples is evaluated separately.

[0048] Preferably, the analytical data comprise an increase in signal resolution as a function of acquisition time. Also preferably, the analytical data comprise spectroscopic, imaging, sensor, or scanning data. More preferably, the data further

comprise measurements made using Raman, luminescence, ultraviolet-visible molecular absorbance, atomic absorbance, infra-red, near infrared, surface plasmon resonance, mass spectrometry, X-ray, nuclear magnetic resonance, refractometry, interferometry, scattering, inductively coupled plasma, atomic force microscopy, scanning tunneling microscopy, microwave evanescent wave microscopy, near-field scanning optical microscopy, atomic fluorescence, laser-induced breakdown spectroscopy, Auger electron spectroscopy, X-ray photoelectron spectroscopy, ultrasonic spectroscopy, dielectric spectroscopy, microwave spectroscopy, resonance-enhanced multiphoton ionization, and the like. Also, combinations, improvements, and subclasses of these techniques can be used, for example surface plasmon resonance and fluorescence, Raman and infrared, and any others. The techniques used to collect the analytical data may also comprise photon probe microscopy, electron probe microscopy, ion probe microscopy, field probe microscopy, scanning probe microscopy, and the like. In an embodiment, analytical data is provided using techniques relying on collection of electromagnetic radiation in the range from 0.05 Angstroms to 500 millimeters (mm).

[0049] The sample may be inorganic material, organic material, polymeric material, biological material, or combinations thereof. In an embodiment the sample comprises polycarbonate. Further, concentration ranges of species of interest analyzed using these techniques can range from detected single molecules to concentrations of up to 100 percent of materials of interest. Any type of single molecule spectroscopy and imaging can utilize the method of this invention.

[0050] In an embodiment, the invention comprises computer readable media comprising software code for performing the methods of the invention. The present invention also includes systems comprising the method of the invention.

[0051] In another aspect, the present invention comprises an apparatus for analyzing at least one sample for a parameter of interest using a pre-determined signal quality response function which comprises: a collecting system for collecting analytical data comprising the parameter of interest from a sample; a processing

system for processing the analytical data; a screening rate accelerator toolbox for applying mathematical transform analysis to the data; a data analysis system for determining whether the data processed by the screening rate accelerator toolbox satisfies a pre-determined signal quality response function value; and a statistical toolbox for analyzing the processed data for the parameter of interest.

[0052]   In an embodiment, the mathematical transform analysis comprises multivariate analysis. Preferably, the multivariate analysis comprises neural networks analysis, principal components analysis, partial least squares analysis, linear multivariate analysis, or nonlinear multivariate analysis.

[0053]   In another embodiment, the mathematical transform analysis comprises discrete transform analysis. Alternatively, the mathematical transform analysis may comprise continuous transform analysis.

[0054]   In yet another embodiment, the mathematical transform analysis comprises time averaging analysis, smoothing analysis or Savitsky-Golay analysis.   The mathematical transform analysis may also comprise Fourier transform, Gabor transform, or Hadamard transform.

[0055]   In an embodiment, the mathematical transform analysis comprises wavelet transform. Preferably, the parameters of the mathematical transform are determined during the course of analysis.

[0056]   In an embodiment, the pre-determined signal quality response function comprises one or more measured signal parameters. In an embodiment, at least one of the measured signal parameters comprises a signal-to-noise ratio.   In another embodiment, at least one of the measured signal parameters comprises signal resolution. In another embodiment, at least one of the measured signal parameters comprises peak shift. In yet another embodiment, at least one of the measured signal parameters comprises signal distortion.

[0057]   In an embodiment, the apparatus of the invention includes at least one energy source for interacting with a sample. Preferably, the energy source is a light

source, an ion source, or a radiation source. In an embodiment, the apparatus does not contain a light source. The data collection system may be an optical spectrometer, an ion spectrometer, a mass detector, an imaging camera, or other instrument capable of quantifying spectral or imaging information.

[0058]   In an embodiment, the analytical data comprises a first-order array. In an alternate embodiment, the analytical data comprises a multi-order array.

[0059]   In an embodiment, evaluation of each individual sample in an array of samples is substantially simultaneous. In an alternate embodiment, each individual sample in an array of samples is evaluated separately.

[0060]   Preferably, the analytical data comprise an increase in signal resolution as a function of acquisition time. Also preferably, the analytical data comprise spectroscopic, imaging, sensor, or scanning data. More preferably, the data further comprise measurements made using Raman, luminescence, ultraviolet-visible molecular absorbance, atomic absorbance, infra-red, near infrared, surface plasmon resonance, mass spectrometry, X-ray, nuclear magnetic resonance, refractometry, interferometry, scattering, inductively coupled plasma, atomic force microscopy, scanning tunneling microscopy, microwave evanescent wave microscopy, near-field scanning optical microscopy, atomic fluorescence, laser-induced breakdown spectroscopy, Auger electron spectroscopy, X-ray photoelectron spectroscopy, ultrasonic spectroscopy, dielectric spectroscopy, microwave spectroscopy, resonance-enhanced multiphoton ionization, and the like. Combinations, improvements, and subclasses of these techniques can also be used. The techniques used to collect analytical data may also comprise photon probe microscopy, electron probe microscopy, ion probe microscopy, field probe microscopy, scanning probe microscopy, and the like. In an embodiment, analytical data is provided using techniques relying on collection of electromagnetic radiation in the range from 0.05 Angstroms to 500 millimeters (mm). For example, detection of thermal or luminescence emission is performed using spectroscopy or imaging. In yet another

embodiment, the apparatus comprises computer readable media comprising software code.

[0061]   The methods and apparatus described herein provide a means to reduce sample integration time for spectroscopic data, imaging arrays, sensor array data and the like. As used herein, integration time is defined as the time during which a detector acquires incoming photons, and data acquisition time is defined as the total time required to collect, process, and store individual signals from an analyzed sample or its certain portion.

[0062]   The method of the present invention may be used to analyze a parameter of interest in a sample, wherein a parameter of interest comprises a chemical, physical, or mechanical aspect of the sample which can be monitored experimentally. Parameters of interest include, but are not limited to, starting reaction components, chemical intermediates, reaction by-products, final products, and mechanical parameters such as moduli, and the like. For example, in the synthesis of melt polycarbonate, a parameter of interest may be the starting components bisphenol A or diphenyl carbonate, oligomer intermediates, phenol, Fries products, or the final polymer.

[0063]   As described herein, the sample may comprise individual samples, multiple individual samples arranged in a fixed format (i.e. multi-element arrays), as well as a plurality of individual samples. Multi-element arrays may be arranged in a geometrically defined array. Thus, in an embodiment, a sample comprises a combinatorial library. In an embodiment, individual regions in the sample array or library are evaluated separately. In an alternate embodiment, evaluation of the entire array or library is substantially simultaneous.

[0064]   The methods and devices of the present invention apply mathematical transform analysis to reduce the integration time required to measure a parameter of interest in a sample. Any method of mathematical transform analysis which results in a decrease in sample integration time as a result of the mathematical transformation of the data may be used. Thus, in an embodiment, the mathematical transform analysis comprises multivariate analysis using multivariate measurement, wherein multivariate

measurement comprises measurements of more than one variable or response for each sample, and multivariate analysis is a mathematical analysis in which more than one variable or response is analyzed for each sample (Beebe, K. R., et al., In Chemometrics: A Practical Guide, p. 6, Wiley, New York, NY, 1998). For example, a whole or a portion of a collected optical spectrum comprising multiple wavelengths may be analyzed for a single analyte (Beebe, K. R., et al., at 6).

[0065]   Types of multivariate analysis include linear multivariate analysis, nonlinear multivariate analysis, partial least squares analysis, principal components analysis, and neural networks analysis. Generally, linear multivariate analysis is used to describe linear relationships between independent and dependent variables using straight line calibration functions, and non-linear multivariate analysis is used to describe nonlinear relationships between independent and dependent variables (Otto, M., In Chemometrics: Statistics and Computer Application in Analytical Chemistry, p. 215, Wiley-VCH, Weinheim, Germany, 1999). Principal components analysis and factor analysis are transforms that find and interpret hidden complex and possibly causally determined relationships between features in a data set; the correlating features are then converted to factors which are themselves non-correlated (Einax, J. W., et al., In Chemometrics in Environmental Analysis, p. 164, VCH, Weinheim, 1997). Partial least squares (PLS) analysis and principal components regression (PCR) analysis make use of an inverse calibration approach where it is possible to calibrate for the desired component(s) while implicitly modeling the other sources of variation; the difference between PLS and PCR is in how the factors are calculated (Beebe, K. R., et al., Chemometrics: A Practical Guide, p. 279, Wiley, New York, NY, 1998). Neural networks analysis describes analysis of a data set to a specific problem by iterative adjustment of weights in a net during the learning process; this adaptation may be done either by comparison of the desired result with the data at the output of the net (supervised learning) or by maximizing differences in the learning data based on an arbitrary criterion of similarity (unsupervised learning) (Hierlemann, A., et al., In Sensors Update, Vol. 2, pp. 119-180, Eds. H. Baltes, W. Göpel, and J. Hesse, VCH, Weinheim, 1996).

[0066]   The mathematical transform analysis may comprise continuous transform analysis or discrete transform analysis. Continuous transform analysis comprises the analysis of continuous signals, such as for example, analogue voltage, which is defined for all values of time (C. L. Phillips and J. M. Parr, In Signals, Systems, and Transforms, p. 2, Prentice Hall, Upper Saddle River, NJ, 1999). An example of continuous transform is integration (C. L. Phillips and J. M. Parr, at 390). Discrete transform analysis comprises the analysis of discrete signals such as the intensity of light reaching each pixel of a diode array detector (C. L. Phillips and J. M. Parr, at 2). An example of discrete transform is summation (C. L. Phillips and J. M. Parr, at 390).

[0067]   The method of mathematical transform may also comprise time averaging analysis where measurements are repeated and the corresponding data added (Vandeginste, B.G.M., et al., Handbook of Chemometrics and Qualimetrics, Part B, p. 538-539, Elsevier, Amsterdam, The Netherlands, 1998) or smoothing analysis which uses an accumulation of values of a number of data points in a small segment or window in the same scan. The smoothing function may be a block function, a polynomial function or an exponential function (Vandeginste, B.G.M., et al., at 535-553). Alternatively, the method of mathematical transform may also comprise Savitsky-Golay analysis which is a combination of smoothing and derivative analysis (Vandeginste, B.G.M., et al., at 550).

[0068]   The method of mathematical transform may also comprise Fourier transform, Gabor transform, or Hadamard transform. Fourier transform comprises a method of representing mathematical models of signals and systems in the frequency domain using sine and cosine functions (C. L. Phillips and J. M. Parr, In Signals, Systems, and Transforms, p. 174, Prentice Hall, Upper Saddle River, NJ, 1999), Gabor transform comprises a short-time Fourier transform employing Gaussian weight functions (T. Masters, Signal and Image Processing With Neural Networks: A C++ Sourcebook, p. 150, Wiley, New York, NY, 1994), and Hadamard comprises encoding signals using periodic square wave functions (P.J. Treado and M.D. Morris, Analytical Chemistry 61: 723A-734A (1989).

[0069] The method of mathematical transform may also comprise wavelet transform. As used herein, the term wavelet transform analysis comprises the use of an inverse transform matrix which applies an infinite set of possible basis functions comprising wavelets, mother wavelets, analyzing wavelets and the like, and wherein the window within the time-frequency plane is allowed to vary.

[0070] The wavelet transform analysis may comprise a wavelet de-noising algorithm, where a wavelet de-noising algorithm is a mathematically defined combination of wavelet transforms which are designed specifically to de-noise spectra. In an embodiment, the adaptive wavelet de-noising algorithm tries many different combinations of wavelet filters and threshold/shrinkage options to maximize the signal quality response function (e.g. signal-to-noise ratio) for that spectrum. After shrinkage, an inverse transform is computed to produce a wavelet processed signal.

[0071] Like Fourier transforms, wavelet transforms comprise an inverse transform matrix that can be viewed as a rotation in function space. These transforms provide a means of mapping signals from one domain to another, such as from the time-scale domain to the frequency domain. Unlike Fourier transform analysis, which is limited to transformation functions utilizing only sines and cosines, wavelet transforms utilize an infinite set of possible basis functions, and allow for the window within the time-frequency plane to vary. The flexibility inherent to wavelet analysis allows analysis of both short basis, high-frequency functions (to isolate signal discontinuities) as well as long basis, low-frequency functions (to obtain detailed frequency analysis) (see e.g., Amara, IEEE Computational Sciences and Engineering, 2:50-61 (1995)).

[0072] Wavelet analysis has been applied for processing signals for a variety of applications including the analysis of spectra. For example wavelet analysis has been used to determine relative ion abundances in mass spectrometry (US Patent 5,436,447), to provide classification of sensor data (US Patent 5,561,431), to evaluate peak intensities in flow-injection analysis (Bos, M. et al., Anal. Chim. Acta, 267:73-80 (1992)), to extract the relevant components from spectral data for multivariate

calibration (Jouan-Rimbaud, D., et al., Anal. Chem., 69:4317-4323 (1997)), for pattern recognition (Walczak, B., et al., Anal. Chem., 68:1742-1747 (1996)), and for classification of spectra (Bos, M., et al., Chemom. Intell. Lab. Syst., 23:115-122 (1994)). Wavelets have been used to compress spectra (Bos, M. et al., Chemom. Intell. Lab. Syst., 23:115-122 (1994); Chau, F. T. et al., Appl. Spectrosc., 50:339-348 (1996)), to standardize spectra (Walczak, B., et al., Chemom. Intell. Lab. Syst., 36:41-51 (1997)), and to characterize non-parameterized magnetic resonance (Young, K., et al., Magn. Reson. Med., 40:816-821 (1998)) and oscillographic signals (Zheng, J., et al., Analyst, 124:893-896 (1999)).

[0073] Wavelets are often used to reduce noise via wavelet shrinkage and thresholding techniques (see e.g. Donoho, D., Different Perspectives on Wavelets, Proceeding of Symposia in Applied Mathematics, 47:173-205 (1993)). Wavelets have been used to remove noise from synthetic Gaussian peaks (Mittermayr, C. R. et al., Chemom. Intell. Lab. Syst., 34:187-202 (1996)), near-IR (Walczak, B., et al., Chemom. Intell. Lab. Syst., 36:81-94 (1997)) and NMR spectra (Walczak, B. et al., Trends Anal. Chem., 16:451-463 (1997)), and images such as SIMS (Nikolov, S. G., et al., Chemom. Intell. Lab. Syst., 34:263-273 (1996)) and ultrasound (US Patent 5,619,998; US Patent 5,497,777; and Alsberg, B. K. et al., Chemom. Intell. Lab. Syst., 37:215-239 (1997)).

[0074] Wavelet filters are a means to apply pre-determined wavelet functions. The wavelet used may include wavelets of the Daubechies family, the Symmlet family, the Haar family, the Coiflet family, the Beylkin family, and the like. Within each family are wavelet subclasses distinguished by the number of coefficients and the level of iteration employed. Wavelets may also be classified in a family by the number of vanishing moments, where vanishing moments comprise a set of mathematical relationships directly related to the number of wavelet coefficients. For example, Daubechies asymmetrical filters are minimal phase filters that generate wavelets that have minimal support for a given number of vanishing moments. Symmlets are also wavelets with a minimum size support for a given number of vanishing moments, but

are generally symmetrical. Both filters are defined by the number of vanishing moments, which differ depending on the shape and size of the analyte signal.

[0075] The de-noising algorithm may also include a thresholding procedure. A threshholding procedure, such as soft thresholding or wavelet shrinking, comprises a mathematical transformation whereby wavelet coefficients caused by noise can be reduced. Thus, threshholding may be applied to the wavelet coefficients to shrink or remove those wavelet coefficients mainly due to noise components. Two key parameters for the shrinkage step are the shrinkage method and the low frequency cutoff. These parameters are preferably optimized for each spectrum using shrinkage methods including, but not limited to, Visu, SURE, and MAD, which are known to those of skill in the art (see e.g. D.L. Donoho and I.M. Johnstone, WAVELAB™ toolbox available at http://www-stat.Stanford.edu/~wavelab).

[0076] The method comprises applying mathematical transform analysis to achieve a preset, or pre-determined, signal quality response function value. The pre-determined signal quality response function, y=f(x), by which the processed spectra are evaluated may vary depending on the nature of the experiment. The inputs to the signal quality response function (y=f(x)) may consist of one or more signal parameters either used individually or combined together mathematically.

[0077] For example, the pre-determined spectral parameter (x) may comprise the signal-to-noise ratio (S/N) at specific peaks in the spectra. As defined herein, signal-to-noise ratio is defined as the amplitude of the measured response for a component of interest in the sample as compared to the root-mean square of the measured response due to general background.

[0078] In another embodiment, the signal quality parameter of interest comprises peak shift in a specific region of the spectra, where peak shift comprises a change in the position (as opposed to the amplitude) of a measured signal. For example, spectroscopic peaks which correspond to known starting materials may shift due to chemical reaction or exposure to a specific environment.

[0079]   In yet another embodiment, the signal quality parameter comprises band distortion, where band distortion comprises a change in the shape of a measured signal.   For example, spectroscopic peaks which correspond to known reaction materials may be distorted due to physical force/stress applied to the sample, or chemical/electronic interaction with the surrounding environment.   Also, un-optimized wavelet de-noising parameters such as the number of vanishing moments or the threshold/shrinkage method may also cause band distortion.   Methods of quantifying band distortion include calculating of peak area, as for example, using integration techniques such as, but not limited to, the extended trapezoidal rule. Alternatively band distortion may be quantified by measuring peak (band) width at either the base or half-height, using methods known to those in the art.

[0080]   In another embodiment, the signal quality parameter comprises band or peak resolution.   As defined herein, band resolution is a measure of how completely two neighboring peaks or bands are separated from each other, wherein increasing resolution comprises increasing band separation, and a resolution of zero indicates two neighboring peaks are completely overlapped with each other.   For example, resolution may be calculated as $R_s = 2 * [(R1-R2)/(W1+W2)]$ for Gaussian-shaped peaks, where R1 and R2 are the locations of the neighboring peaks, and W1 and W2 are the widths at half-height for the neighboring peaks, or other methods standard in the art (D. A. Skoog and G. G. Leary, Principals of Instrumental Analysis, 4th Ed., Saunders College Publishing, Fort Worth, TX, 1992, p. 592).

[0081]   The method may be used to analyze data presented as first-order or higher order arrays. A first order array comprises a one dimensional data set, such as, but not limited to, spectra, chromatograms, ionograms, sensor array data, and the like. A second-order array comprises data sets which are two-dimensional, such as images of combinatorial libraries, gas chromatograph-mass spectroscopy data (GC/MS), excitation-emission fluorescence maps, time dependent monitoring of processes, and the like.

[0082]   The present invention also includes systems comprising the method of the invention.   The system can be a stand-alone system that performs the analysis of samples directly, or it can be incorporated in a more general system that also includes a separation step, followed by a detection step.   The separation can be performed using any system that analyzes relatively large amounts of materials or a system that analyzes very small amounts of materials (nanogram, femtogram, and less).   An example of the latter system can be a lab-on-a-chip system.   As another example, a system may comprise a sensor element followed by a detection step.   As described herein, sensors generally include a material that changes its spectroscopic or other property as a function of analyte concentration in proximity to this element.

[0083]   The invention also includes an apparatus for analyzing at least one sample for a parameter of interest using a pre-determined signal quality response function which comprises: (a) a collecting system for collecting analytical data comprising a parameter of interest from a sample; (b) a processing system for processing the data; (c) a screening rate accelerator toolbox for applying wavelet transform analysis to the data; (d) a data analysis system for determining whether the wavelet processed data satisfies a pre-determined signal quality response function value; and (e) a statistical toolbox for analyzing the processed data for the parameter of interest.

[0084]   As used herein, a collecting system is defined as a set of hardware and software components to collect spectroscopic or imaging signals.   The hardware components can include any components needed to generate and record the signals from the sample region of interest.   A processing system is defined as a set of hardware and software components to process the collected analytical data into a form acceptable for further analysis, and a screening rate accelerator toolbox is essentially a collection of mathematical functions used to apply mathematical transform analysis to the data in an interactive or iterative manner.   As used herein, a data analysis system is defined as a system that compares measured data from a sample (e.g. the signal quality response function value) to a pre-determined standard (e.g. a pre-determined signal quality response function value).   The measured data may be processed by the screening rate accelerator toolbox or unprocessed.   Finally, a statistical toolbox is

essentially a collection of mathematical functions that can be used analyze data (wavelet processed or unprocessed) for a parameter of interest.

[0085] The samples may be analyzed using sensor array techniques. As used herein, a sensor array is a set of sensor elements combined with a single or multiple detectors. Each sensor element can include a material that changes its spectroscopic or other property as a function of analyte concentration in proximity to the element. Using sensor array techniques, a spectroscopically inactive (undetectable) analyte can be detected with a spectroscopic or imaging system that utilizes the method of the invention.

[0086] In an embodiment, the apparatus of the invention includes no light source for exciting a sample. In this case, detection of thermal or luminescence emission is performed using spectroscopy or imaging. Luminescence emission can include chemoluminescence, bioluminescence, triboluminescence, electroluminescence, and any other type of radiation emission generated by a process that does not involve an absorption of incoming photons and a process that does include absorption of incoming photons.

[0087] In an embodiment, the invention describes the use of mathematical transform analysis to reduce the data acquisition time for spectroscopic analysis of sample arrays. The present invention is particularly useful for the analysis of signal from each element of a multi-element sample system such as a combinatorial library and the like. One benefit of the screening rate accelerator toolbox of the present invention can be expressed as the reduction of integration time ΔT needed to achieve a predetermined signal response function, such as a defined signal-to-noise ratio (S/N), in a collected signal before, T_b, and after, T_a, applying the screening rate accelerator toolbox:

$$\Delta T = T_b - T_a$$

[0088] Increases in screening throughput can potentially be achieved by using certain hardware capabilities. For example, an increase in the power of a light source

in optical spectroscopic applications can be used for the reduction of integration time of spectral and/or image acquisition. However, an increase in excitation power often causes premature photo-bleaching and other types of sample degradation. Another potential approach for decreasing integration time is by reduction in the noise level of a photodetector. In most systems, however, such noise reduction requires significant cooling of the detector, which adds to the size, cost, and complexity of the system.

[0089]   In an embodiment, the invention uses a screening rate accelerator toolbox comprising wavelet transform to reduce data acquisition time. For example, the reduction in data acquisition time can be accomplished by applying wavelet algorithms to de-noise data sets. The de-noised data allows for a preset signal quality response function (e.g. S/N = 20) to be obtained with a reduced sample integration time, which thereby reduces the time required for data acquisition for each sample. In an embodiment, the data sets are one-dimensional, such as spectra, chromatograms, ionograms, sensor array data and the like. In another embodiment, the data sets are two-dimensional, such as images of combinatorial libraries, gas chromatograph-mass spectroscopy data (GC/MS), excitation-emission fluorescence maps, time-dependent monitoring of processes, and the like. Preferably, reduction in signal integration results in a significant acceleration in the screening process.

[0090]   Referring now to Figure 1, in one aspect, the invention comprises a method for reducing the time required for analyzing a sample for a parameter of interest which comprises: (10) selecting a pre-determined signal quality response function; (20) selecting a pre-determined integration time T_a; (30) collecting analytical data from the sample with integration time T_a; (50) applying a screening rate accelerator toolbox comprising mathematical transform analysis to the data, wherein the mathematical transform analysis is performed using conditions designed to achieve a pre-determined signal quality response function value comprising the value obtained when samples are analyzed without mathematical transform analysis using integration time T_b, wherein T_b is greater than T_a; and (90) analyzing the data processed by the screening rate accelerator toolbox for the parameter of interest. In an embodiment, the method also comprises (40) determining whether the analytical data collected from the sample

with integration time T_a satisfies the pre-determined signal quality response function value prior to applying the screening rate accelerator toolbox and applying the screening rate accelerator toolbox if the collected data does not satisfy the pre-determined signal quality response function value, but not if the collected data does satisfy the pre-determined signal quality response function value. Also, in an embodiment, if the collected analytical data does not require application of the screening rate accelerator toolbox to satisfy a pre-determined signal quality response function value, a shorter pre-determined integration time T_a is selected, and the method performed using the new value for T_a (45).

[0091] The method permits iterative application of the screening rate accelerator toolbox. Thus, in an embodiment, the method includes (60) determining whether the data processed by the screening rate accelerator toolbox satisfies the pre-determined signal quality response function value; and if the processed data does not satisfy the pre-determined signal quality response value, (70) re-applying the screening rate accelerator toolbox using a mathematical transform analysis different from the analysis previously applied until the data processed using the screen rate accelerator toolbox either satisfies the pre-determined signal quality response function value or comprises an optimized signal quality response function value.

[0092] In some cases, even with application of the screening rate accelerator toolbox, the data collected with an integration time T_a may not satisfy the signal quality response function value. Thus, in an embodiment, the method includes (80) repeating the method with a new, larger value for T_a if the data which is optimized using the screening rate accelerator toolbox does not satisfy the pre-determined signal quality response function value.

[0093] Generally, the method requires selection of a parameter of interest and an evaluation of the limits of the data resolution for the parameter of interest. Referring now to Figure 2, in the high throughput screening (HTS) method of the invention, a parameter of interest, such as a specific spectral peak which corresponds to a distinct chemical component (e.g. diphenyl carbonate), or sample parameter (e.g. molecular

weight) is chosen (11) to be measured using a preferred analytical method (12). The analytical method may comprise spectral or imaging techniques known to those in the art, such as Raman spectroscopy, fluorescence spectroscopy, absorbance spectroscopy, and the like. A specific signal quality response function (e.g., the signal-to-noise ratio, S/N, for a specific spectral region) is then chosen (13). Alternatively, the inputs to the signal quality response function may consist of multiple signal parameters (e.g. S/N and peak resolution) combined together mathematically ($y = f(x_1, x_2, \ldots x_n)$). For example, a multi-criteria signal quality response function can be evaluated by summation of individual signal quality parameters multiplied by weighting factors ($y = w_1{}^*x_1 + w_2{}^*x_2 + \ldots w_n{}^*x_n$), where $w_1, w_2 \ldots w_n$ are specific weighting factors, and $x_1, x_2 \ldots x_n$ are the outputs of individual signal quality parameters. It will be recognized that other common mathematical manipulations such as logarithms, exponentials, cross-products, and the like, can be used to combine the individual signal quality parameters together mathematically to form a signal quality response function.

[0094] Using reference data measured for controls comprising preset values for the parameter of interest (21), the screening rate accelerator toolbox of the present invention is applied to the data (22) and a minimum integration time, (T_a), which satisfies the preset signal quality response function value (y) is chosen (23). For example, reference spectra or data may be used to select baseline values comprising an absence of analyte, and thereby define noise (N), and to find spectral regions which are specific to the reaction components of interest, and thereby define signal (S). An integration time which allows for a pre-determined signal-to-noise ratio (e.g. of 20, or any other pre-determined value) at a specific peak or set of peaks is then determined using the reference data. This integration time is then used for the subsequent screening of samples comprising unknown amounts of the parameter(s) of interest.

[0095] Figure 3 depicts the method of the present invention using wavelet transform as the mathematical transform of the screening rate accelerator toolbox for the analysis of experimental sample 102 which is one of several elements arranged in sample array 104. For example, a typical sample array comprises a 384-well

microtiter plate reactor, or similar array. The parameter of interest is measured (30), using the appropriate pre-determined spectral or imaging technique (e.g. Raman spectroscopy). The data may then be evaluated (40) to determine whether it satisfies a pre-determined signal quality response function value or whether to apply the screening rate accelerator toolbox of the invention (50). Alternatively, the data may automatically be processed by the screening rate accelerator toolbox without pre-analysis step (40). The screening rate accelerator toolbox allows for iterative application of wavelet transforming algorithms to the dataset of interest to transform the data and thereby reduce sample integration time and provide an accelerated screening rate. In an embodiment, the algorithm de-noises the data. Preferably, the algorithm comprises wavelet filters (51). Also preferably, the algorithm comprises threshold/shrinkage parameters (52). In an embodiment, the parameters of the wavelet transformation algorithm are determined during the course of analysis.

[0096] Preferably, there is at least one wavelet algorithm for de-noising the collected data to achieve the preset signal quality response function value (e.g. a pre-determined signal-to-noise ratio) at a reduced integration time. In an embodiment, the adaptive wavelet de-noising method of the invention allows for the choice of wavelet filter and thresholding/shrinking parameters to be optimized in real-time using criteria specified beforehand. Thus, criteria may be specified based on values obtained using reference spectra of known components. Alternatively, wavelet de-noising may be accomplished by optimizing the wavelet filter to obtain the required wavelet coefficients. For example, software can be programmed to search through a bank of available filters and options until a filter is found which meets some pre-determined signal quality response function (e.g., a signal to noise ratio, S/N, of 100).

[0097] After shrinkage, an inverse transform is preferably computed to produce a wavelet processed signal (53). The wavelet processed signal can then be tested to determine signal quality, as for example, by computing the improved S/N (54).

[0098] The next step (60) comprises an evaluation of whether the processed data satisfies the pre-determined signal quality response function value (e.g., S/N=20). In

an embodiment, this step may be by-passed and samples analyzed for the parameter of interest (90). In an embodiment, samples for which wavelet processed signals that do not meet the pre-set signal quality response function value may be processed using a different algorithm characterized by a different combination of parameter settings or variation (decrease or increase) in the integration time (70). Preferably, the parameters of the wavelet transform algorithm are determined during the course of analysis. Thus, in an embodiment, the results from the (n) de-noising are used to evaluate parameters for the (n+1) de-noising, and results from the (n) and (n+1) de-noising are used to evaluate parameters for the (n+2) de-noising, and so on. In an embodiment, parameter settings, such as wavelet filter, number of vanishing moments, shrinkage method, low frequency cutoff, that produce the highest quality spectrum are chosen as optimal may be used as starting points for the analysis of other samples.

[0099] For the wavelet processed data that does meet the pre-set signal quality response function, the signal parameter of interest is quantified and stored for processing, chemometric analysis, and selection of those samples which appear to comprise the best candidates based on the reaction being assessed (90).

[0100] In an embodiment, each sample of a larger array is analyzed individually. Alternatively, samples comprising a multiple arrays (or multi-order arrays) (103) may be processed simultaneously (Figure 3). For example, an entire multi-element array can be analyzed using imaging techniques such as, but not limited to: (a) the analysis of scattered, reflected, or transmitted light from samples irradiated with light from a light source; (b) the analysis of radiation emission from samples upon excitation with a source of electromagnetic or ionizing radiation; and (c) the analysis of thermal or luminescence radiation emitted from samples during a chemical reaction, mechanical testing, and the like. The array can be analyzed using sensor array techniques where each sample in the array is analyzed with an individual sensor or a sensor array and the data is collected from the sensors simultaneously.

[0101] In one aspect, the invention comprises an apparatus for analyzing a sample for at least one pre-selected spectroscopic or imaging property, which comprises: (a) a collecting system for collecting analytical data comprising the parameter of interest from the sample; (b) a processing system for processing the imaging or spectral data; (c) a screening rate accelerator toolbox for applying mathematical transform analysis to the data; (d) a data analysis system for determining whether the wavelet processed data satisfies the pre-determined signal quality response function; and (e) a statistical toolbox for analyzing the processed data for the parameter of interest.

[0102] Thus, referring now to Figures 4 and 5, the apparatus may comprise a sample array 104 where at least one well comprising individual sample 102 is analyzed using collecting system 126. In an embodiment, the sample 102 is irradiated with excitation light 106 generated from light source 108. Excitation light 106 can be filtered via wavelength selection component 110 prior to transmission via coupling optics component 112 and illuminating fiber optic cable 114 to probe 116 for irradiation of sample 102. Similarly, emitted radiation 118 collected by probe 116 passes via collecting fiber optics cable 120 through coupling optics component 122 and emission wavelength selection component 124 prior to processing by processing system 128. The signal is then transmitted for transformation using screening rate accelerator toolbox 130 and then processed by data analysis system 132 for selection of signals that satisfy a pre-determined signal quality response function value prior to processing by statistical toolbox 134 for analysis of the de-noised data for a parameter of interest.

[0103] In an embodiment, the apparatus is suited for analysis of individual sample wells (Figure 4). In an embodiment, and referring now to Figure 5, the apparatus is suited for simultaneous analysis of a multi-element array 104. For simultaneous analysis of a multi-element array 104, a single or multiple light source 108, with associated wavelength selectors 110 may be positioned at selected angles to illuminate an array 104 of samples with excitation light 106. Emitted radiation 118 is then transmitted to imaging detector 128, preferably after processing by wavelength selection component 124. The signal is then transmitted for wavelet processing by the

screening rate accelerator toolbox 130 and then processed by data analysis system 132 for selection of signals that satisfy a pre-determined signal quality response function criteria prior to processing by statistical toolbox 134.

[0104] In some embodiments, the apparatus of the invention may not include a light source. For example, a light source is not required for measurement of luminescence emission or thermal emission from a sample or set of samples. Luminescence may occur during a chemical reaction, or as a result of a mechanical perturbation as in for example a performance test of solid material.

[0105] The screening rate accelerator toolbox may be applied to analysis of Raman or other spectral analysis. For example, Raman bands at 911 cm^{-1} and 1072 cm^{-1} can be used for the quantification of diphenyl carbonate (DPC), a starting component used in melt polymerization synthesis of polycarbonate. For Raman spectroscopic analysis of small-scale polycarbonate samples, however, detector integration time for unprocessed (original) spectra typically ranges from tens of seconds to several minutes for each sample. This can result in data acquisition times of several hours for even small (<1,000) combinatorial libraries of samples.

[0106] For example, Figures 6 and 7 illustrate the signal-to-noise (S/N) improvements that result upon applying a screening rate accelerator toolbox comprising wavelet transform to de-noise Raman spectra of diphenyl carbonate (DPC), wherein the dashed lines correspond to the original spectra and the solid lines correspond to wavelet processed spectra, located near the 1072 cm^{-1} and 911 cm^{-1} peaks, respectively. In an embodiment, signal to noise enhancement may be determined by dividing the S/N of the spectra after de-noising by the S/N ratio of the original spectra.

[0107] The increased S/N resulting from the adaptive wavelet transform method of the invention can be used to decrease integration time. For example, referring now to Figure 8, an original Raman spectra for the diphenyl carbonate 1072 cm^{-1} band 144 achieves a maximal S/N value 140 at an integration time 142 of about 50 to 900 seconds. For the wavelet transformed spectra 146, this same S/N is obtained after an

integration time 142 of 5 seconds, thus allowing for a relative improvement in integration time (T_b/T_a) by at least factor of 10 and a ΔT of at least 45 seconds per sample (horizontal arrow). Referring now to Figure 9, for untransformed the 911 cm^{-1} diphenyl carbonate band 150, the maximum S/N value 140 is obtained after an integration time 142 of 900 seconds. For the wavelet transformed spectra 152 this S/N value 140 is obtained after an integration time 142 of 10 seconds, thereby resulting in a relative improvement in integration time (T_b/T_a) by about 90-fold and a ΔT of about 890 seconds per sample (horizontal arrow).

[0108] In an embodiment, the reduction in integration time ΔT is a function of the pre-determined signal quality response function of interest. Where S/N is the signal quality response function of interest, the reduction in integration time ΔT generally increases as the S/N function value is increased. The increase may be linear (Figure 10) or may plateau at higher S/N values (Figure 11). Thus, for measurements requiring highly stringent criteria (e.g. a high S/N), the method of the present invention may save significant amounts of integration time per sample.

[0109] The reduction in signal integration time upon applying the screening rate accelerator toolbox of the present invention correlates to a overall reduction in data acquisition time. In an embodiment, the screening rate accelerator toolbox of the invention can dramatically reduce integration time per sample, and is particularly useful for libraries having multiple samples, each requiring integration times exceeding about 10 msec. Figure 12 shows the amount of time saved for screening combinatorial libraries ranging from 10 to 5,000 diphenyl carbonate samples when the screening rate accelerator toolbox of the invention is applied.

[0110] Other types of spectroscopic measurements are suitable for the methods of the present invention. Also, the sample may be inorganic material, organic material, polymeric material, biological material, or combinations thereof. For example, the screening rate accelerator toolbox of the present invention has been applied to reduce sample integration time for light scattering measurements at 500 nanometers (nm) of combinatorial libraries of protective organic hard coatings. In an embodiment,

wavelet processing by the method of the invention de-noises the spectra so that a predetermined S/N ratio (S/N=25) is found with a lower integration time for wavelet processed measurements (i.e. 50 msec) as opposed to non-processed measurements (i.e. 500 msec), thus, resulting in a 20-fold reduction in integration time, and an overall savings in time of 450 msec/sample. Time savings on a msec scale can be important for monitoring of rapid dynamic processes important in such fields as medicine, rapid reaction kinetics, physics, biology, and the like.

[0111] As will be recognized by those of ordinary skill in the art, all or part of the steps in the method of the present invention may be coded or otherwise written in computer software, in a variety of computer languages including, but not limited to, C, C++, Pascal, Fortran, Visual Basic, Microsoft Excel, MATLAB, Mathematica, Java, and the like. Accordingly, additional aspects of the present invention include computer software for performing one or more of the method steps set forth herein. The software code may be compiled and stored in executable form on computer readable media as, for example, computer ROM, floppy disk, optical disk, hard disks, CD ROM, or the like. The invention may be further understood by reference to the following non-limiting examples.

Example 1

[0112] Raman spectroscopy can be used for the high throughput screening of combinatorial libraries. Integration time for Raman spectroscopy, however, is dependent on sample size. For samples ranging from 10 mL to 0.01 mL (the size used for combinatorial analysis), integration time typically ranges from tens of seconds to several minutes. Thus, to achieve the high throughput of multiple samples, it is important to decrease the integration time required for acquisition of data while maintaining a S/N which allows for quantification of the reaction components of interest.

[0113] The production of polycarbonate by melt polymerization utilizes diphenyl carbonate (DPC) and bis-phenol A (BPA). Monitoring the ratio of these two starting

31

components can be useful for optimizing the overall reaction. In this example, Raman bands at 911 cm^{-1} and 1072 cm^{-1} were used for quantification of sample DPC.

[0114] Measurements were performed using a fiber-optic Raman system that consisted of a spectrograph, a laser, and a fiber-optic probe. The Raman spectrograph (Model Echell-NIR775; EIC Raman Systems, Norwood, MA) covered a spectral range from 200 to 3500 cm^{-1} with a 4 cm^{-1} resolution. This capability was available by using a gold-coated echelle grating (52.65 lines/mm) that dispersed light in two dimensions to fully exploit the CCD detector area. The detector was a cryogenically cooled CCD camera (Model CH270, Photometrics Inc.; Tuscon, AZ). The light source was a wavelength-stabilized high-power diode laser (300 mW output power, 785 nm emission wavelength, Model SDL-8530; SDL Inc., San Jose, CA). Raman spectra were collected using a data acquisition package (EIC Raman Systems; Norwood, MA) and converted into ASCII format using GRAMS/32 software (Galactic Industries, Inc., Salem, NH). Raman spectra were collected using a standard 5 m long fiber-optic sampling probe (EIC Raman Systems; Norwood, MA).

[0115] A sample of DPC was measured at 13 different integration times (0.1, 1, 5, 10, 30, 60, 120, 180, 240, 300, 420, 600, and 900 seconds) resulting in a total of 13 spectra for analysis. A reference spectrum of the chemical system under study (in this case DPC) was used to select baseline regions where no analyte exists and to find spectral regions specific to DPC. Bands at 911 and 1072 cm^{-1} bands were selected. Thus, for measurements of DPC samples, spectra were truncated to a total of 1024 points covering the spectral region between 800 cm^{-1} and 1200 cm^{-1}.

[0116] The signal quality response function, $y = f(x)$, used in this example consisted of the signal-to-noise ratio (S/N), which was determined by dividing the signal (intensity at 911 or 1072 cm^{-1}) by the noise (root-mean square intensity between 930 and 950 cm^{-1} for the 911 cm^{-1} band and between 1090 and 1140 for the 1072 cm^{-1} band).

[0117] Adaptive wavelet de-noising algorithms using multiple different combinations of Daubechies and Symmlet wavelet filters and threshold/shrinkage

options were used to maximize S/N for each spectrum. Wavelet de-noising was performed in the MATLAB™ programming environment (Mathworks, Inc., Natick, MA, Version 5.3) using functions written internally and from the Stanford University WAVELAB™ toolbox (http://www-stat.stanford.edu/~wavelab, Version 8.02). For the 13 spectra analyzed in this example, a majority of the optimized wavelets used four vanishing moments for the Daubechies filter, and eight vanishing moments for the Symmlet filters. It is to be understood that in an alternative embodiment, software which searches through a bank of available filters and options until one or combination of filters/options is found which meets a pre-determined signal quality function (e.g. y=S/N of 20) can be used for selection of the appropriate algorithm parameters.

[0118] A threshholding procedure was then applied to the wavelet coefficients to reduce or remove those wavelet coefficients which were mainly due to noise. In this example, soft thresholding or wavelet shrinking methods which reduce the value of wavelet coefficients caused by noise were optimized by varying both the shrinkage method, and the low frequency cutoff each spectrum. Generally, shrinkage methods such as 'Visu', 'SURE', and 'MAD' with low-frequency cutoffs between 2 and 6 were found to be satisfactory for this application (see e.g., D.L. Donoho and I.M. Johnstone, the WAVELAB™ toolbox at http://www-stat.Stanford.edu/~wavelab).

[0119] After application of each shrinkage method, an inverse transform was used to produce a wavelet processed spectrum. The spectra were then tested to determine their signal quality response value, y. The parameter settings (i.e., wavelet filter, number of vanishing moments, shrinkage method, and low frequency cutoff) that produced the highest quality spectrum (in this example, highest S/N) were chosen as optimal. The optimization procedure was repeated for each of the 13 Raman spectra collected at the different integration times.

[0120] Figures 6 and 7 illustrate the S/N improvements that result upon applying the screening rate accelerator toolbox to Raman spectra of melt polycarbonate at the 1072 cm-1 and 911 cm-1 bands, respectively. In these figures, the dashed lines are the

original spectra, and the wavelet processed spectra are shown as the solid lines. The dashed vertical line indicates the position of the peak of interest used for determination of S/N. Signal-to-noise enhancement was determined by dividing the S/N of the spectra after wavelet processing by the S/N ratio of the original spectra. The improvement in S/N comprises a factor of 13.2 for the 1072 cm^{-1} band (Figure 6), and a factor of 2.8 for the 911 cm^{-1} band (Figure 7).

[0121] Figures 8 and 9 show the signal-to-noise (S/N) ratios as a function of integration time for the 1072 cm^{-1} and 911 cm^{-1} bands before (solid circles) and after (open squares) wavelet processing of the data. The increased S/N resulting from the adaptive wavelet de-noising method of the invention can be used to decrease integration time. For the original, unfiltered spectra, the maximal S/N for the 1072 cm^{-1} band was obtained after a 50 to 900 second integration time (Figure 8). For the wavelet processed spectra, this same or higher level of S/N was obtained at integration times of 5 seconds, thus resulting in an at least 10 to 180 fold decrease in integration time. For the 911 cm^{-1} band, the original spectra had a maximum S/N at 900 second integration time. For the wavelet processed spectra, an equivalent S/N value was obtained after 10 second integration time, thus resulting in a 9 fold decrease in integration time (Figure 9).

[0122] As shown in figures 10 and 11, it was found that there is a significant reduction of signal integration time at each S/N when the screening rate acceleration toolbox of the invention is applied to the spectra. The plots shown in these two figures were generated (from figures 8 and 9, respectively) by subtracting the value of the integration time for the original (unprocessed) spectra from the integration time upon applying the screening rate accelerator toolbox for a given signal-to-noise ratio (S/N) for the corresponding spectra. Notably, there was a significant increase for the reduction in the integration time ΔT with an increase in S/N (Figures 10 and 11). For example, for the 1072 cm^{-1} band, to achieve a S/N ratio of 20 required about 147 seconds integration time for the unprocessed spectra, but only 3 seconds for the wavelet processed data, thereby resulting in a savings of about 144 seconds of

integration time per sample (Figure 10). For the 911 cm^{-1} band, the time saved for a S/N=20 is on the order of 10 seconds per sample (Figure 11).

Example 2

[0123]   Figure 12 depicts the total time required to screen libraries of different size before, and after, application of the screening rate accelerator toolbox of the invention. In this example, a preset S/N value of 20 was selected as the S/N value required to achieve necessary precision and accuracy of analysis for the 1072 cm^{-1} DPC Raman band. It was found that to achieve a S/N value of 20, the unprocessed (original) spectra had to be integrated for 147 seconds (about 2.5 min). Using the screening rate accelerator toolbox of the present invention, the same quality of spectrum (S/N=20) is achieved with an integration time of 3 seconds (a 50-fold decrease). This savings of 144 seconds per sample provides a total time savings per library as shown, where the open circles correspond to the total analysis time for a given library of original spectra before, and the closed circles correspond to the total analysis time for a given library after, application of the screening rate accelerator toolbox of the invention.

[0124]   It will be recognized by those in the art that advantages of wavelet transform analysis to reduce data acquisition time disclosed here over other methods for the screening of potential reaction conditions include:

1.   An increase in screening throughput of 10-200 times using exclusively software tools;

2.   A method for rapid, real-time assessment of multiple reactions;

3.   Safe, non-invasive measurements;

4.   Measurement of small reaction volumes;

5.   High-throughput, rapid analysis capable of automation for combinatorial chemistry or production-scale applications;

6.   Analysis of gaseous, dissolved, solid or molten samples;

7.   Application to a large variety of spectroscopic and imaging techniques;

8.   Samples may be inorganic, organic, polymeric or biological; and

9.   Measurements may detect from a single molecule to components making up 100% of the sample.

[0125]   It will be understood that each of the elements described above, or two or more together, may also find utility in applications differing from the types described herein. While the invention has been illustrated and described as a method for rapid, high-throughput, nondestructive analysis of combinatorial reactions, it is not intended to be limited to the details shown, since various modifications and substitutions can be made without departing in any way from the spirit of the present invention. For example, robotics equipment can be used to prepare the samples, and various types of parallel analytical screening methods can be incorporated. Also, other parameters besides Raman analysis of diphenyl carbonate and light scattering measurements of sample hazing can be measured. As such, further modifications and equivalents of the invention herein disclosed may occur to persons skilled in the art using no more than routine experimentation, and all such modifications and equivalents are believed to be within the spirit and scope of the invention as defined by the following claims.